Testing for Unit Roots in Nonlinear Dynamic Heterogeneous Panels

Changli He* and Rickard Sandberg†

Department of Economic Statistics, Stockholm School of Economics, P. O. Box 6501, SE-113 83 Stockholm, Sweden

SSE/EFI Working Paper Series in Economics and Finance No. 582

January 2005

Abstract

In this paper we present a unit root test against a nonlinear dynamic heterogeneous panel with each country modelled as an LSTAR model. All parameters are viewed as country specific. We allow for serially correlated residuals over time and heterogeneous variance among cross sections. The test is derived under three special cases: (i) the number of cross sections and observations over time are fixed, (ii) observations over time are fixed and the number of cross sections tend to infinity, and (iii) first letting the number of observations over time tend to infinity and thereafter the number of cross sections. Small sample properties of the test show modest size distortions and satisfactory power being superior to the Im, Pesaran and Shin t-type of test. We also show clear improvements in power compared to a univariate unit root test allowing for non-linearities under the alternative hypothesis.

JEL classification: C12; C23; C52

Key words: Dynamic Nonlinear Heterogeneous Panels; Structural Breaks; Unit roots; t - statistics; Central limit theorem

---

*e-mail:changli.he@hhs.se

†e-mail:strs@hhs.se

Acknowledgements: This research has been supported by the Jan Wallander's and Tom Hedelius' Foundation, Grant No. J02-35. The paper has been presented at the 11th Biennial Panel Data Conference at College Station, Texas 2004.

1 Introduction

Testing the unit root hypothesis in linear panels has received much attention in recent years. One reason for this is that many of the classical unit root tests in univariate settings suffer from low power against near unit root alternatives, and a remedy to this problem is to consider a panel of several univariate time series. For instance, Quah (1994) considers a completely homogeneous linear panel with no cross section specific effects. This approach is, however, quite unrealistic and is likely to yield biased estimators, see e.g. Hsiao (1986). A more general panel is introduced in Levin, Lin, and Chu (2002) because it allows for cross section specific intercepts and time trends, as well as serially correlated residuals over time and heterogeneous variances among cross sections. Both Quah (1994) and Levin, Lin, and Chu (2002) base their tests of a common unit root on the least squares pooled estimator, and inference is enhanced by letting both the number of cross sections and observations over time tend to infinity, implying, of course, a limited practical use. However, Harris and Tzavalis (1999) study the same issues as in Levin, Lin, and Chu (2002) but they instead derive analytical results under the assumption of a traditional panel set-up, i.e. the time dimension is fixed and the number of cross sections is seen as large.

One drawback with all these tests mentioned above is that their alternative hypotheses imply that e.g. all cross sections converge to a long-run equilibrium at the same rate. This is too strong to be held in any interesting empirical cases, as concluded in Maddala and Wu (1999). The problem with a too strong/unrealistic alternative hypothesis is relaxed by Im, Pesaran, and Shin (2003). They assume under the alternative hypothesis that a fraction of the total number of cross sections possesses a linear mean-reversion possibly with cross section specific convergence rates, and that the remaining cross sections are non-stationary. In addition, their testing procedure is fundamentally different because they are averaging individual unit root t-test statistics.

Much of the ongoing research focuses on generalizing the panel unit root tests as such that the tests allow for dependence among cross sections, see for instance Phillips and Sul (2003) and Peseran (2003), but with the linearity property kept. We will stress another issue because evidence against the linearity of the adjustment process in univariate time series has recently been found, see Leybourne, Newbold, and Vougas (1998), Harvey and Mills (2002), and Lanne, Lütkepohl, and Saikkonen (2003), among others. Our concern is therefore that if a cross section is modelled as nonlinear, the conventional linear panel approach yields unit root tests with modest power, as outlined in He and Sandberg (2005c), and studying a panel is not necessarily a solution to obtaining satisfactory power. Despite the fact that one obtains more observations, the nonlinearity may be too hard to detect. Our aim is therefore to derive a test of a common unit root in a nonlinear dynamic heterogeneous panel allowing for linear, nonlinear, and non-stationary models under the alternative hypothesis. The panel in this paper is a generalization of the panel smooth autoregressive introduced in He and Sandberg (2005c). We are using the testing methodology by Im, Pesaran, and Shin (2003), whereas the testing methodology in He and Sandberg (2005c) is inspired by Harris and Tzavalis (1999).

The unit root tests in this paper are very simple to conduct and the limiting distributions are (mostly) the standard normal. We allow for serially correlated errors over time and heterogeneous variances among cross sections. The following dimensions of the panel are considered: (i) The number of cross sections and observations over time are fixed. (ii) The observations over time are fixed and the number of cross sections tends to infinity. (iii) The number of observations over time and cross sections tend to infinity (sequential limits).

The rest of the paper is organized as follows. In Section 2 we present a nonlinear heterogeneous dynamic panel. Testing procedures of a unit root in nonlinear panels are discussed in Section 3. The unit root tests are presented in Section 4. Asymptotic and finite-sample properties are investigated through Monte Carlo experiments in Section 5. Section 6 concludes. Thereafter an appendix follows with tables and further simulation results.

2 A nonlinear heterogeneous dynamic panel

Consider a sample of n cross sections observed over T time periods (not necessarily the same for each cross section so we allow for unbalanced panels). Suppose that the stochastic process $\{y_{it}\}$ is generated by the first-order panel smooth autoregressive (PSTAR(1)) model

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\pi}_{i1} + \mathbf{x}_{it}'\boldsymbol{\pi}_{i2}F_i(t) + u_{it}, \quad i = 1, ..., n, \quad t = 1, ..., T, \tag{1}$$

where $\mathbf{x}_{it} = (1, y_{i,t-1})'$, $\boldsymbol{\pi}_{i1} = (\pi_{i10}, \pi_{i11})'$, $\boldsymbol{\pi}_{i2} = (\pi_{i20}, \pi_{i21})'$, and $F_i(t)$ is chosen to be the logistic smooth transition function defined by

$$F_i(t; \gamma_i, c_i) = \frac{1}{1 + \exp[-\gamma_i(t - c_i)]} - 1/2. \tag{2}$$

In (2), $\gamma_i \in [0, \infty)$ measures the speed of transition over time from one regime to another, and where $c_i \in (0, T)$ gives the point in time where the transition will be symmetric around. By the model specification in (1) and (2), each cross section unit is modelled as the LSTAR(1) model proposed in Lin and Teräsvirta (1994) (however, modified since we adjust the transition function by subtracting one half). We especially note that all parameters are defined to be heterogeneous for each cross section, by which it also follows that the model in (1) with (2) is a generalization of the panel logistic smooth autoregressive (PLSTAR) model in He and Sandberg (2005c). Furthermore, for $\gamma_i \in [0, \infty)$, the transition function $F_i(t, \gamma_i, c_i)$ is an non-decreasing function in t, and $\gamma_i = 0$ implies, by construction, that $F_i(t; 0, c_i) = 0$ holds for all t, and the resulting model in (1) is a linear panel autoregressive (PAR) model with parameter vector $\boldsymbol{\pi}_{i1}$. On the other hand, letting $\gamma_i \to \infty$, $F_i(t; \infty, c_i)$ becomes an indicator function such that $F_i(t; \infty, c_i) = -0.5$ for $0 \leq t < c_i$ and $F_i(t; \infty, c_i) = 0.5$ for $c_i \leq t < T$. It is seen that the PLSTAR(1) model nests the panel threshold autoregressive (PTAR) model with parameter vectors $\boldsymbol{\pi}_{i1} - 0.5\boldsymbol{\pi}_{i2}$ and $\boldsymbol{\pi}_{i1} + 0.5\boldsymbol{\pi}_{i2}$ in regime one ($0 \leq t < c_i$) and two ($c_i \leq t < T$), respectively. Finally, u_{it} is an error term such that for all i and t, $\{u_{it}\}$ defines a sequence of independently distributed random variables with zero means and heterogeneous variances σ_i^2.

3 Testing procedures

3.1 The test statistic

Testing the unit root hypothesis in the PLSTAR(1) model (1) is achieved by imposing the following parameter restrictions

$$H_0 : \pi_{i10} = 0, \quad \pi_{i11} = 1, \quad \gamma_i = 0, \tag{3}$$

for all i, and the maintained model for each cross section is therefore

$$y_{it} = y_{i,t-1} + u_{it}, \tag{4}$$

i.e. a random walk without drift (this is the most relevant null hypothesis since the PLSTAR(1) model does not contain any time trend). The alternative hypothesis could be described as not H_0, meaning that the alternative hypothesis can be a mixture of LSTAR models (and AR models) and unit root processes. It is, however, important that the total number of LSTAR (and AR models) models, n^*, satisfies $n \geq n^* > 0$ where $n^* = \delta n$ and $\delta \in (0, 1]$, to guarantee consistent tests. Furthermore, as pointed out in Luukkonen, Saikkonen, and Teräsvirta (1988), imposing the restriction $\gamma_i = 0$ leads to identification problem since the parameters π_{i20}, π_{i21}, and c_i are not identified under the null hypothesis. To remedy this problem, we replace the transition function in (2) with its first-order Taylor expansion around $\gamma_i = 0$. This approximation is feasible since $F_i(t)$ is twice differentiable in γ_i and the first derivative evaluated at $\gamma_i = 0$ is non-zero. The approximation yields $F_i(t) \approx 0.25\gamma(t - c_i)$ (ignoring the remainder). Substituting for this approximation into (1) and collecting terms yields the linearized version of the PLSTAR(1) model

$$y_{it} = \tilde{\mathbf{x}}_{it}' \boldsymbol{\alpha}_i + \tilde{u}_{it}, \tag{5}$$

where $\tilde{\mathbf{x}}_{it} = (1, t, y_{i,t-1}, ty_{i,t-1})'$, $\boldsymbol{\alpha}_i = (\alpha_{i1}, ..., \alpha_{i4})'$, and $\tilde{u}_{it}$ is an error term adjusted with respect to the Taylor approximation such that $\tilde{u}_{it} = u_{it}$ holds under the null hypothesis.[1] The transformed null hypothesis of a unit root is now given by

$$H_0^{aux} : \alpha_{i1} = 0, \quad \alpha_{i2} = 0, \quad \alpha_{i3} = 1, \quad \alpha_{i4} = 0, \tag{6}$$

for all i. We proceed by examining the single hypothesis $\alpha_{i3} = 1$, by running the regression in (5) for each cross section and calculating the t_i statistic given by

$$t_i = \frac{\hat{\alpha}_{i3} - 1}{\hat{\sigma}_{\hat{\alpha}_{i3}}}, \quad i = 1, ..., n, \tag{7}$$

where $\hat{\alpha}_{i3}$ denotes the OLS estimator of α_{i3}, $\hat{\sigma}_{\hat{\alpha}_{i3}} = S_i^2 \mathbf{r}_1 \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}' \tilde{\mathbf{x}}_{it} \right)^{-1} \mathbf{r}_1'$, $S_i^2 = \sum_{t=1}^{T} (y_{it} - \tilde{\mathbf{x}}_{it}' \hat{\boldsymbol{\alpha}}_i)^2 / (T-4)$, and $\mathbf{r}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$. However, to simplify matters in following finite-sample analysis, we focus on a modified t-statistic, denoted t_i^m, which is defined by

$$t_i^m \equiv \frac{\hat{\alpha}_{i3} - 1}{\tilde{\sigma}_{\hat{\alpha}_{i3}}}, \quad i = 1, ..., n, \tag{8}$$

where $\tilde{\sigma}^2_{\hat{\alpha}_{i3}}$ is defined as in (7) but S_i^2 is replaced with

$$\tilde{S}_i^2 = \Delta \mathbf{y}_i' \mathbf{M} \Delta \mathbf{y}_i / (T-1),$$

where $\Delta \mathbf{y}_i = (\Delta y_{i1}, ..., \Delta y_{iT})'$, $\mathbf{M} = \mathbf{I}_T - \boldsymbol{\iota}_T \boldsymbol{\iota}_T' / T$ is the within transformation matrix where $\boldsymbol{\iota}_T$ is a $T \times 1$ vector of ones. It is clear that both S_i^2 and $\tilde{S}_i^2$ converge under the null hypothesis in probability to σ_i^2, and thus, the asymptotic distribution for t_i and t_i^m is the same. However, the finite-samples properties for S_i^2 and $\tilde{S}_i^2$ differ.[2] Moreover, our choice of $\tilde{S}_i^2$ is arbitrary and is based on that $\mathbf{M}$ is fixed (non-stochastic) and generates consistent estimates of σ_i^2. Other options, however not analyzed here, would be $\mathbf{M} = \mathbf{I}_T$ or $\mathbf{M} = \mathbf{I}_T - \mathbf{X}_T (\mathbf{X}_T' \mathbf{X}_T)^{-1} \mathbf{X}_T'$ where $\mathbf{X}_T = (\boldsymbol{\iota}_T, \boldsymbol{\tau}_T)$ and $\boldsymbol{\tau}_T = \{1, ..., T\}$. Before stating how the information from each cross section through the t_i^m test statistic in (8) is used to conduct a test for a common unit root, we establish some fundamental finite-sample properties for t_i^m under the null hypothesis (6).

3.2 Finite-sample properties of the test statistic under the null hypothesis

Three important finite-sample properties of the t_i^m test statistic under the null hypothesis can be observed. The two first properties concern the invariance with

[1] This holds because the remainder from the approximation equates to zero under the null hypothesis.

[2] Some finite sample properties for S_i^2 and $\tilde{S}_i^2$ are established in Table 8 in the Appendix A.

respect to y_{i0} and the heterogeneous variances σ_i^2. The last property is that, for all i, and T sufficiently large, the second moment of t_i^m exists. One way to readily confirm the first two properties is to vectorize the auxiliary model in (5) and divide the regressors into stochastic and deterministic matrices according to

$$\mathbf{y}_i = \mathbf{X}_T \boldsymbol{\alpha}_{i1} + \mathbf{Z}_T \boldsymbol{\alpha}_{i2} + \tilde{\mathbf{u}}_i, \tag{9}$$

where $\mathbf{X}_T$ is defined as before, $\boldsymbol{\alpha}_{i1} = (\alpha_{i1}, \alpha_{i2})'$, $\mathbf{Z}_T = (\mathbf{y}_{i,-1},\ \mathbf{D}_T \mathbf{y}_{i,-1})$, $\mathbf{y}_{i,-1} = (y_{i0},\ y_{i1}, ..., y_{i,T-1})'$, $\mathbf{D}_T = diag\left\{\boldsymbol{\tau}_T\right\}$ is a $T \times T$ matrix with a time trend on its diagonal, $\boldsymbol{\alpha}_{i2} = (\alpha_{i3}, \alpha_{i4})'$, and $\tilde{\mathbf{u}}_i = (\tilde{u}_{i1}, ..., \tilde{u}_{i,T})'$. The partitioned regression in (9) implies that we can rewrite (8) under the null hypothesis (6) as

$$t_i^m = \frac{\mathbf{r}_2 \left[(\mathbf{Q}_T \mathbf{Z}_T)' (\mathbf{Q}_T \mathbf{Z}_T) \right]^{-1} \times \left[(\mathbf{Q}_T \mathbf{Z}_T)' (\mathbf{Q}_T \mathbf{u}_i) \right]}{\tilde{S}_i \sqrt{\mathbf{r}_2 \left[(\mathbf{Q}_T \mathbf{Z}_T)' (\mathbf{Q}_T \mathbf{Z}_T) \right]^{-1} \mathbf{r}_2'}}, \tag{10}$$

where $\mathbf{r}_2 = \left[\begin{array}{cc} 1 & 0 \end{array}\right]$ and $\mathbf{Q}_T = \mathbf{I}_{_T} - \mathbf{X}_T \left(\mathbf{X}_T' \mathbf{X}_T\right)^{-1} \mathbf{X}_T'$. Furthermore, under (6) we can express $\mathbf{y}_{i,-1}$ as

$$\mathbf{y}_{i,-1} = \boldsymbol{\iota}_T y_{i0} + \mathbf{C}_T \mathbf{u}_i, \tag{11}$$

where

$$\mathbf{C}_T = \left[\begin{array}{cccccc} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{array}\right]_{T \times T}. \tag{12}$$

Using (11) we obtain that $\mathbf{Z}_T = (\boldsymbol{\iota}_T y_{i0} + \mathbf{C}_T \mathbf{u}_i,\ \mathbf{D}_T \boldsymbol{\iota}_T y_{i0} + \mathbf{D}_T \mathbf{C}_T \mathbf{u}_i)$, and because $\mathbf{Q}_T$ is orthogonal to both $\boldsymbol{\iota}_T y_{i0}$ and $\mathbf{D}_T \boldsymbol{\iota}_T y_{i0}$, $\mathbf{Q}_T \mathbf{Z}_T = (\mathbf{Q}_T \mathbf{C}_T \mathbf{u}_i, \mathbf{Q}_T \mathbf{D}_T \mathbf{C}_T \mathbf{u}_i)$ holds. Furthermore, under the null hypothesis $\Delta \mathbf{y}_i = \mathbf{u}_i$ which yields $\tilde{S}_i^2 = \mathbf{u}_i' \mathbf{M} \mathbf{u}_i / (T-1)$. It follows, for all i and T, that the t_i^m test statistic is invariant with respect to y_{i0}

To verify the second property, write t_i^m as

$$\begin{aligned} t_i^m &= \sqrt{T-1} \times \frac{\mathbf{r}_2 \left[(\mathbf{Q}_T \mathbf{Z}_T / \sigma_i)' (\mathbf{Q}_T \mathbf{Z}_T / \sigma_i) \right]^{-1}}{\sqrt{(\mathbf{M} \mathbf{u}_i / \sigma_i)' (\mathbf{M} \mathbf{u}_i / \sigma_i)}} \\ &\quad \times \frac{\left[(\mathbf{Q}_T \mathbf{Z}_T / \sigma_i)' (\mathbf{Q}_T \mathbf{u}_i / \sigma_i) \right]}{\sqrt{\mathbf{r}_2 \left[(\mathbf{Q}_T \mathbf{Z}_T / \sigma_i)' (\mathbf{Q}_T \mathbf{Z}_T / \sigma_i) \right]^{-1} \mathbf{r}_2'}}. \end{aligned} \tag{13}$$

Obviously the σ_i's cancel out in (13), and in the matrix $\mathbf{Q}_T \mathbf{Z}_T / \sigma_i = (\mathbf{Q}_T \mathbf{C}_T \mathbf{u}_i / \sigma_i, \mathbf{Q}_T \mathbf{D}_T \mathbf{C}_T \mathbf{u}_i / \sigma_i)$ we see that the sub-vectors $\mathbf{u}_i / \sigma_i$ are identically distributed with mean $E[\mathbf{u}_i / \sigma_i] = \mathbf{0}$ and covariance matrix $E[\mathbf{u}_i \mathbf{u}_i' / \sigma_i^2] = \mathbf{I}_{_T}$. We conclude that t_i^m is invariant with respect to the nuisance parameter σ_i.

The third property, i.e. that $E\left(t_i^m\right)^2$ exists, is confirmed by simulations. It appears that $E\left(t_i^m\right)^2$ exists for $T \in [4, \infty]$, also indicating that the second moment is existing in an asymptotic sense. In a related context Im, Pesaran, and Shin (2003) also rely upon simulations to conclude that the second moment of their test statistic exists. The present case is somewhat more complicated and the proof of this is left for further research. The methods in Larsson (1997) and Nabeya (1999) might be applicable for $T \in [4, \infty)$ and $T \to \infty$, respectively.

The finite-sample properties for the t_i^m tests statistic summarized above combined with that t_i^m defines a measurable function of the i.i.d. sequence $\{\mathbf{u}_i/\sigma_i\}$, are sufficient conditions for us to conduct tests of a common unit root in the nonlinear heterogeneous panel in (1).

4 The unit root tests

By imposing certain restrictions on n and T we obtain many interesting testing situations in the nonlinear dynamic heterogeneous panel described above.

4.1 Fixed T unit root test in a balanced panel letting n tend to infinity

Assumption 1 *Let $\{u_{it}\}$ be a sequence of independently and normally distributed random variables such that, for all i and t, $E[u_{it}] = 0$ and $E[u_{it}^2] = \sigma_i^2 \in \mathbb{R}_{++}$ hold.*

Assumption 2 *Let T be the same for all i (a balanced panel).*

Proposition 1 *Under Assumptions 1 and 2, the null hypothesis (6), and $3 < T < \infty$, the individual statistics, t_i^m, are i.i.d. with $E[t_i^m] = \mu(T) \in \mathbb{R}$ and $V[t_i^m] = \eta^2(T) \in \mathbb{R}_{++}$. Furthermore, define the random variable $\mathcal{Z}_0 \equiv n^{-1/2} \sum_{i=1}^{N} \left(t_i^m - \mu(T)\right) / \eta(T)$. Then, by letting n tend to infinity, the Lindberg-Lèvy central limit theorem gives*

$$\mathcal{Z}_0 \xrightarrow{d}_n N(0,1). \tag{14}$$

where $\xrightarrow{d}_n$ denotes convergence in distribution by letting $n \to \infty$.

In Proposition 1 it is clear that we use the information from each cross section unit by averaging the t_i^m test statistics, inspired by Im, Pesaran, and Shin (2003). This procedure is in contrast to the one in He and Sandberg (2005c) where data is pooled and a test for a common unit root is based on the LSDV estimator. It should be pointed out that both the mean $E[t_i^m]$ and the variance $V[t_i^m]$ are under the null hypothesis functions of T and are tabulated in Table 8 in the Appendix A for different values of T.

4.2 Fixed T unit root test in an unbalanced panel letting n tend to infinity

Assumption 3 *Let T be different for at least one i (an unbalanced panel).*

Proposition 2 *Under Assumptions 1 and 3, the null hypothesis (6), and $3 < T < \infty$, the individual statistics, t_i^m, are independently heterogeneously distributed with $E[t_i^m] = \mu_i(T) \in \mathbb{R}$, $V[t_i^m] = \eta_i^2(T) \in \mathbb{R}_{++}$, and for any i, $E[(t_i^m)^3] < \infty$, holds. Furthermore, define the random variable $\mathcal{Z}_1 \equiv n^{-1/2}(\bar{t} - \bar{\mu}(T)) / (\bar{\eta}(T))$ where $\bar{t} = n^{-1}\sum_{i=1}^n t_i^m$, $\bar{\mu}(T) = n^{-1}\sum_{i=1}^n \mu_i(T)$, and $\bar{\eta}(T) = n^{-1}\sum_{i=1}^n \eta_i^2(T)$. Then, by letting n tend to infinity, the Liapounov central limit theorem gives*

$$\mathcal{Z}_1 \xrightarrow{d}_n N(0,1). \tag{15}$$

In Proposition 2 we must assert that $E[(t_i^m)^3] < \infty$ which holds as long as $3 < T$, and is confirmed simulations, however not reported here.

4.3 Fixed T and n unit root test in a balanced panel

This is perhaps the most relevant case from a practitioner point of view. However, viewing the number of cross sections as finite ruins the asymptotic inference concluded in Propositions 1 and 2. For both a fixed T and n, the analytical expression for the finite-sample distribution of $n^{-1}\sum_{i=1}^n t_i^m$ is hardly known. However, in Subsection 3.2 it is shown that the t_i^m test statistic is nuisance parameter free under the null hypothesis. This means that the finite-sample distribution for $n^{-1}\sum_{i=1}^n t_i^m$ is readily obtained by Monte Carlo simulations for any combinations of T and n. These simulations are reported in the next section, and they also provide a measure of how well the n finite-sample distributions approximate the asymptotic $N(0,1)$ distribution.

4.4 Asymptotic T and n unit root test

We apply the method of sequential limits, see Phillips and Moon (1999).[3] Specifically, we consider first a fixed cross section unit and let T tend to infinity which yields intermediate asymptotic results, and thereafter we let n tend to infinity. The intermediate univariate asymptotic results that we need follow from Lemma 4 in He and Sandberg (2005a).

[3] Lemma 6 in Phillips and Moon (1999) might be applicable to find conditions for sequential convergence to imply joint convergence.

Proposition 3 *If Assumption 1 and the null hypothesis in (6) hold, then for any i and letting T tend to infinity it follows that*

$$t_i^m \xrightarrow{d}_T \xi_i = \frac{\mathbf{r}_1 \mathbf{\Psi}_{1i}^{-1} \mathbf{\Pi}_{1i}}{\left(\mathbf{r}_1 \mathbf{\Psi}_{1i}^{-1} \mathbf{r}_1'\right)^{1/2}}, \tag{16}$$

where $\xrightarrow{d}_T$ denotes convergence in distribution by letting $T \to \infty$, and

$$\mathbf{\Psi}_{1i} = \begin{bmatrix} \mathbf{M}_{11i} & \mathbf{M}_{12i} \\ \mathbf{M}_{12i} & \mathbf{M}_{13i} \end{bmatrix}, \quad \mathbf{\Pi}_{1i} = \begin{bmatrix} \mathbf{\Pi}_{11i} \\ \mathbf{\Pi}_{12i} \end{bmatrix},$$

with sub-matrices given by

$$\mathbf{M}_{11i} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix},$$

$$\begin{aligned}
\mathbf{M}_{12i} &= \begin{bmatrix} \int_0^1 B_i(r)dr & \int_0^1 rB_i(r)dr \\ \int_0^1 rB_i(r)dr & \int_0^1 r^2 B_i(r)dr \end{bmatrix}, \\
\mathbf{M}_{13i} &= \begin{bmatrix} \int_0^1 B_i^2(r)dr & \int_0^1 rB_i^2(r) \\ \int_0^1 rB_i^2(r)dr & \int_0^1 r^2 B_i^2(r)dr \end{bmatrix}, \\
\mathbf{\Pi}_{11i} &= \begin{bmatrix} B_i(1) \\ B_i(1) - \int_0^1 B_i(r)dr \end{bmatrix}, \\
\mathbf{\Pi}_{12i} &= \begin{bmatrix} 0.5\left(B_i(1)^2 - 1\right) \\ 0.5\left(B_i(1)^2 - \int_0^1 B_i^2(r)dr - 1/2\right) \end{bmatrix},
\end{aligned}$$

where $B_i(r)$ denotes a standard Brownian motion with respect to u_i on $[0,1]$. Assume further that, $B_i(r)$ and $B_j(r)$ are independent for $i \neq j$. For $i = 1, ..., n$, the limiting distributions ξ_i are i.i.d. with $E[\xi_i] = \mu^* \in \mathbb{R}$ and $V[\xi_i] = \eta^{*2} \in \mathbb{R}_{++}$ where $\lim_{T\to\infty} \mu(T) = \mu^*$ and $\lim_{T\to\infty} \eta^2(T) = \eta^{2*}$. Furthermore, define the random variable $\mathcal{Z}_2 \equiv n^{-1/2} \sum_{i=1}^n (\xi_i - \mu^*)/\eta^*$. Then by letting n tend to infinity, the Lindberg-Lèvy central limit theorem gives,

$$\mathcal{Z}_2 \xrightarrow{d}_{n,T} N(0,1), \tag{17}$$

where $\xrightarrow{d}_{n,T}$ denotes convergence in distribution by first letting $T \to \infty$ followed by $n \to \infty$.

In Table 8 in the Appendix A it is seen that $\mu^* = -1$ and $\eta^{2*} = 1.432$.

4.5 Asymptotic T and n unit root test in a panel with serially correlated errors

From the results in He and Sandberg (2005b) we note that a unit root test that accommodates serially correlated errors over time can easily be imposed by adding the terms $\triangle y_{i,t-j}$ ($j \geq 1$) to the auxiliary regression model in (5). This principle is the analogy to the classical ADF tests. We obtain

$$y_{it} = \tilde{\mathbf{x}}_{it}' \boldsymbol{\alpha}_i^a + \triangle \mathbf{y}_{it}' \boldsymbol{\zeta}_i + u_{it}, \tag{18}$$

where $\tilde{\mathbf{x}}_{it}$ is the same vector of explanatory variables as in (5), $\boldsymbol{\alpha}_i^a = (\beta_i, \delta_i, \rho_i, \psi_i)'$, $\triangle \mathbf{y}_{it} = (\triangle y_{i,t-1}, \ \triangle y_{i,t-2}, ..., \triangle y_{i,t-p_i+1})'$, $p_i \geq 2$ denotes the order of augmentation for cross section i, $\boldsymbol{\zeta}_i = \left(\zeta_{i1}, \zeta_{i2}, ..., \zeta_{i,t-p_i-1}\right)'$, and u_{it} is an error term that fulfills Assumption 1.[4]

Assumption 4 *Assume that, for all i, the roots of the characteristic polynomial $\phi(z) = 1 - \zeta_{i1} z - \zeta_{i2} z^2 - ... - \zeta_{i,p_i-1} z^{p_i-1}$ lie outside the unit circle.*

Assumption 4 rules out the possibility of a cross section unit being integrated of order two. The auxiliary null hypothesis of a single unit root for each cross section, under Assumption 4, can now be formulated as

$$H_0^{aux} : \beta_i = 0, \quad \delta_i = 0, \quad \rho_i = 1, \quad \psi_i = 0, \quad \boldsymbol{\zeta}_i \in \mathbb{R}^{p_i-1}, \tag{19}$$

for all i. The alternative hypothesis is not H_0^{aux}, now meaning a mixture of LSTAR models (and AR models) and unit root processes with serially correlated errors, though still in such proportions that the test remains consistent. Under the null hypothesis in (19) the following data generating process (DGP) is obtained

$$y_{it} = y_{i,t-1} + \zeta_{i1} \triangle y_{i,t-1} + \zeta_{i2} \triangle y_{i,t-2} + ... + \zeta_{i,p_i-1} \triangle y_{i,t-p_i+1} + u_{it}, \tag{20}$$

and under Assumption 4 we obtain

$$\Delta y_t = \varepsilon_{it}, \tag{21}$$

where $\varepsilon_{it} = (1 - \zeta_{i1} L - \zeta_{i2} L^2 - ... - \zeta_{i,p_i-1} L^{p_i-1})^{-1} u_{it} = \Psi(L) u_{it}$, where L denotes the lag operator and $\Psi(L)$ is a one-sided moving average polynomial in the lag operator. The process in (21) clearly defines a stochastic process with serially correlated increments. Moreover, proceed by running the regression in (18) to calculate, for each i, the augmented t_i statistic, denoted t_i^a, by

$$t_i^a \equiv \frac{\hat{\rho}_i - 1}{\hat{\sigma}_{\hat{\rho}_i}}, \quad i = 1, ..., n, \tag{22}$$

[4] The regression in (18) corresponds to the NPADF testing equations in He and Sandberg (2005b).

where $\hat{\rho}_i$ denotes the OLS estimator of ρ_i, $\hat{\sigma}_{\hat{\rho}_i} = S_{ia}^2 \left(\mathbf{r}_3 \left[\sum_{t=1}^{T} \tilde{\mathbf{X}}_{it}' \tilde{\mathbf{X}}_{it} \right]^{-1} \mathbf{r}_3' \right)$, $S_{ia}^2 = \sum_{t=1}^{T} \left(y_{it} - \tilde{\mathbf{x}}_{it}' \hat{\boldsymbol{\alpha}}_{i1}^a - \triangle \mathbf{y}_{it}' \hat{\boldsymbol{\zeta}}_i \right)^2 / (T - 4 - p_i + 1)$, $\mathbf{r}_3 = \begin{bmatrix} \mathbf{r}_1' & \mathbf{0}_{1\times(p_i-1)}' \end{bmatrix}$, and $\tilde{\mathbf{X}}_{it} = \begin{bmatrix} \tilde{\mathbf{x}}_{it}' & \triangle \mathbf{y}_{it}' \end{bmatrix}'$. Notice that in the case of serial correlation and for a finite T, one encounters the problem with nuisance parameters. For a finite T and any i, the t_i^a test statistic is dependent on σ_i^2, $\{\zeta_{ij}\}_{j=1}^{p_i-1}$, p_i, and the starting values $\mathbf{y}_{i0} = (y_{i,-p_i+1}, ..., y_{i0})'$. However, if one assume that $\mathbf{y}_{i0} = \mathbf{0}$, then t_i^a still depends upon $\{\zeta_{ij}\}_{j=1}^{p_i-1}$and p_i, but the invariance with respect to σ_i^2 is resurrected.[5] As a result, letting $\mathbf{y}_{i0} = \mathbf{0}$, implies that the expected value and the variance of t_i^a only will depend on p_i, $\{\zeta_{ij}\}_{j=1}^{p_i-1}$, and T, and which is denoted $E[t_i^a(T, p_i, \boldsymbol{\zeta}_i)]$ and $V[t_i^a(T, p_i, \boldsymbol{\zeta}_i)]$ respectively. In Appendix A these expected values and variances are tabulated for different values of T, p_i, and $\boldsymbol{\zeta}_i$. However, the factors that t_i^a is dependent upon in finite-samples are eliminated if one let $T \to \infty$. We conclude the following important result.

Proposition 4 *If Assumptions 1 and 4 and the null hypothesis (19) hold, then, for any i and letting T tend to infinity, it follows that*

$$t_i^a \xrightarrow{d}_T \xi_i, \tag{23}$$

where ξ_i is the same limiting distribution as in Proposition 3. Therefore, by defining the random variable $\mathcal{Z}_3 \equiv n^{-1/2} \sum_{i=1}^{n} (\xi_i - \mu^*) / \eta^*$ and letting n tend to infinity, we obtain the same result as in (17).

Proof. The proof is similar to the proof of Corollary 2.7 in He and Sandberg (2005b), and is therefore omitted. ∎

5 Monte Carlo experiments

In this section we examine the finite-sample properties of the $\mathcal{Z}_0$ and $\mathcal{Z}_3$ test statistics in Proposition 1 and Proposition 4 respectively, by Monte Carlo simulations . The two first Monte Carlo experiments assess the size properties of the tests. In the remaining Monte Carlo experiments, the empirical power of the tests are examined. All experiments are carried under the assumption of balanced panels.

[5]For an example, let $p_i = 2$ in (18), and pre-multiply the matrix of stochastic regressors implied by (18) $[\boldsymbol{\iota}_T y_{i0} + \mathbf{C}_T \mathbf{u}_i, \mathbf{D}_T \boldsymbol{\iota}_T y_{i0} + \mathbf{D}_T \mathbf{C}_T \mathbf{u}_i, \mathbf{c}_0 + \mathbf{C}_T \mathbf{u}_i]$ where $\mathbf{c}_0 = (y_{i0} - y_{i,-1}) \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}'$ with the $\mathbf{Q}_T$ matrix. Notice that $\mathbf{Q}_T$ is orthogonal to $\boldsymbol{\iota}_T y_{i0}$ and $\mathbf{D}_T \boldsymbol{\iota}_T y_{i0}$ but not to $\mathbf{c}_0$. This indicates that t_i^a is dependent on the two starting values y_{i0} and $y_{i,-1}$. Furthermore, recall that σ_i cancels out in (13), but in the the present case division with σ_i yields expressions on the form $\mathbf{c}_0/\sigma_i$ that will not further simplify, and σ_i will affect the test statistic through the starting values.

5.1 Size properties

5.1.1 Estimated size in the case of no serial correlation

The first Monte Carlo experiment examines the $\mathcal{Z}_0$ test statistic (no serial correlation) in Proposition 1 and its size properties when the DGP is given by

$$y_{it} = y_{i,t-1} + u_{it}, \quad t = 1, ..., T, \quad i = 1, ..., n. \tag{24}$$

In (24), it is assumed that $u_{it} \sim nid(0, \sigma_i^2)$ and $\sigma_i^2 \sim U[0.5, 1.5]$ for all i and where U denotes the uniform distribution. The size results are shown in Table 1.

Table 1: The size of the $\mathcal{Z}_0$ test in Proposition 1. No serial correlation.

n \ T	25	50	100	250
5	0.051 (0.101)	0.050 (0.100)	0.049 (0.101)	0.050 (0.099)
25	0.049 (0.099)	0.050 (0.101)	0.049 (0.100)	0.050 (0.100)
50	0.048 (0.098)	0.051 (0.100)	0.051 (0.100)	0.050 (0.101)

Notes: The nominal sizes of the test are 5% and 10% (in parentheses). The results are based on 10 000 replications.

We see in Table 1 that the size distortions for all T and n of the $\mathcal{Z}_0$ test are negligible at both 5% and 10% significance levels. We conclude that the convergence to the asymptotic standard normal distribution is very fast. Already for n as small as five, the approximation to the standard normal distribution is excellent.

5.1.2 Estimated size in the case of serially correlated errors

In the second Monte Carlo experiment we examine the size properties of the $\mathcal{Z}_3$ test statistic (serially correlated errors) in Proposition 4 when u_{it} in (24) is replaced by the AR(1) process, $u_{it} = \rho u_{it} + v_{it}$, where $\rho \in \{0.3, 0.6\}$ and $v_{it} \sim nid(0, 1)$ for all i and t. These results are presented in Table 2.

In Table 2 with $p_i = 0$, the $\mathcal{Z}_3$ test is the same test as $\mathcal{Z}_0$, and is undersized for all values of T, n, and ρ considered. This is to be expected since a test statistic which does not take the serial correlation into account is used. With $p_i = 1$, a first attempt to adjust for the serial correlation, we see that the size distortions are reduced for all T, but increase slowly with n. This supports the well-known fact that over-fitting (the case $p_i \geq 1$), is less harmful than under-fitting (the case $p_i < 1$). We also see that the distortions are more severe when ρ is increased. Moreover, for $T > 50$, $p_i \geq 1$, and all n, the size distortions are modest, however always larger for a more persistent autocorrelation, and that the size distortions

Table 2: The size of the $\mathcal{Z}_3$ test in Proposition 4. Serial correlation.

n	ρ	p_i	0	1	2	3
			$T = 25$			
5	0.3		0.01 (0.02)	0.05 (0.09)	0.05 (0.10)	0.06 (0.12)
	0.6		0.00 (0.00)	0.05 (0.09)	0.04 (0.10)	0.05 (0.11)
25	0.3		0.00 (0.00)	0.04 (0.07)	0.04 (0.08)	0.06 (0.11)
	0.6		0.00 (0.00)	0.03 (0.06)	0.02 (0.05)	0.02 (0.05)
50	0.3		0.00 (0.00)	0.03 (0.06)	0.03 (0.07)	0.06 (0.11)
	0.6		0.00 (0.00)	0.02 (0.03)	0.01 (0.03)	0.01 (0.03)
			$T = 50$			
n	ρ	p_i	0	1	2	3
5	0.3		0.00 (0.00)	0.05 (0.10)	0.05 (0.11)	0.06 (0.12)
	0.6		0.00 (0.00)	0.05 (0.10)	0.05 (0.10)	0.05 (0.10)
25	0.3		0.00 (0.00)	0.04 (0.09)	0.05 (0.09)	0.06 (0.12)
	0.6		0.00 (0.00)	0.03 (0.07)	0.03 (0.07)	0.03 (0.06)
50	0.3		0.00 (0.00)	0.04 (0.08)	0.05 (0.10)	0.06 (0.13)
	0.6		0.00 (0.00)	0.02 (0.05)	0.02 (0.05)	0.02 (0.05)
			$T = 100$			
n	ρ	p_i	0	1	2	3
5	0.3		0.00 (0.01)	0.05 (0.10)	0.05 (0.10)	0.05 (0.10)
	0.6		0.00 (0.00)	0.05 (0.10)	0.05 (0.10)	0.05 (0.10)
25	0.3		0.00 (0.00)	0.04 (0.09)	0.05 (0.10)	0.05 (0.09)
	0.6		0.00 (0.00)	0.04 (0.07)	0.04 (0.08)	0.04 (0.08)
50	0.3		0.00 (0.00)	0.04 (0.08)	0.05 (0.10)	0.04 (0.08)
	0.6		0.00 (0.00)	0.03 (0.06)	0.03 (0.07)	0.03 (0.07)
			$T = 250$			
n	ρ	p_i	0	1	2	3
5	0.3		0.00 (0.01)	0.04 (0.09)	0.05 (0.10)	0.05 (0.10)
	0.6		0.00 (0.01)	0.05 (0.10)	0.05 (0.10)	0.05 (0.10)
25	0.3		0.00 (0.00)	0.04 (0.08)	0.04 (0.09)	0.05 (0.10)
	0.6		0.00 (0.00)	0.04 (0.09)	0.05 (0.09)	0.04 (0.09)
50	0.3		0.00 (0.00)	0.04 (0.08)	0.04 (0.09)	0.05 (0.10)
	0.6		0.00 (0.00)	0.04 (0.08)	0.04 (0.09)	0.04 (0.09)

Notes: The nominal sizes of the test are 5% and 10% (in parentheses). The results are based on 10 000 replications.

are more evident when n is large relatively to T. To this end it is noticed that the remarks about the size for the $\mathcal{Z}_3$ test are in line with the findings for the the IPS test statistic reported in Im, Pesaran, and Shin (2003).

5.2 Empirical power

5.2.1 A heterogeneous nonlinear panel

In the third Monte Carlo study we examine the empirical power of the $\mathcal{Z}_0$ test statistic in Proposition 1 when the DGP accommodates smooth heterogeneous shift in levels and dynamics and is given by

$$y_{it} = \pi_{i10} + \pi_{i11} y_{i,t-1} + (\pi_{i20} + \pi_{i21} y_{i,t-1})\, \tilde{F}_i(t) + u_{it}, \tag{25}$$

and where we have replaced, without loss of generality, the transition function in (2) with $\tilde{F}_i(t) = 1/(1 + \exp[-\gamma_i(t - c_i)])$. This yields more convenient interpretation of the parameters because $\tilde{F}_i(t)$ has the range $[0, 1]$. The parameters in the PLSTAR(1) in (25) are assigned the following values

$$\begin{aligned} &\pi_{i10} = 0, && \pi_{i11} \sim U[0.35, 0.45], && \gamma_i \sim U[0.5, 1.5], \\ &c_i \sim U[0.4T, 0.6T], && \pi_{i20} \sim U[0.5, 1.5], && \pi_{i21} \sim U[0.4, 0.5], \\ &u_{it} \sim nid(0, 1). \end{aligned} \tag{26}$$

The choice of parameter values in (26) implies that all the cross section units display time series that start from the same level because $\pi_{i10} = 0$ holds for all i. The level of a new long-run cross section specific equilibrium is given by $\pi_{i20}/(1 - \pi_{i11} - \pi_{i21}) \in [2, 30]$ (assuming that a complete transition takes place). The parameters π_{i20} are set to vary modestly because it is well known that a test based on a first-order Taylor approximation is not designed to capture changes in the intercept, see Luukkonen, Saikkonen, and Teräsvirta (1988) and the discussion in He and Sandberg (2005a) about the level leverage effect. The panel autoregressive parameters in the linear part in the PLSTAR(1) model yield PAR(1) processes that are modestly persistent because $\pi_{i11} \sim U[0.35, 0.45]$. The panel autoregressive parameters in the nonlinear part of the PLSTAR(1) model are chosen such that $\max_i \pi_{i11} + \max_i \pi_{i21} = 0.95$ (< 1) and the trajectories are therefore stable around the new long-run equilibrium.[6] The speed of transition between regimes varies in the cross sections from 0.5 to 1.5, and implies that a complete transition takes place for all sample sizes that will be considered. The timing of the

[6] We use the word stable rather than stationary, since the transition function is a function of time. Imposing the restriction $\max_i \pi_{i11} + \max_i \pi_{i21} = 0.95 < 1$ rules out the case of a unit root at the end of the sample period.

transitions varies around the mid-point of the sample, i.e. $c_i = 0.5T$, which for instance illustrates that some cross sections respond earlier to, say, a shock and the responses for the other cross sections are somewhat delayed. Typical trajectories generated by (25) and (26) is depicted in Figure 1.

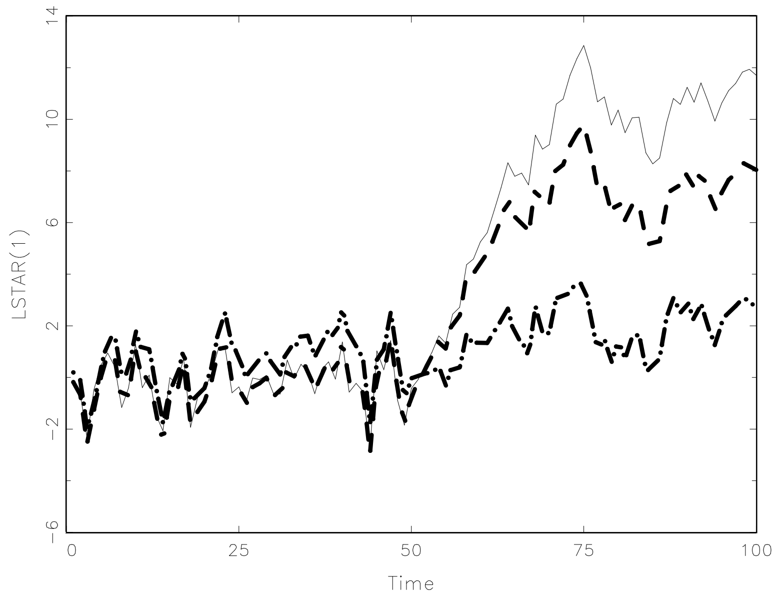

Figure 1: Typical realizations of the LSTAR(1) models in the third (dashed line), fourth (dot-dashed line), and fifth (solid line) Monte Carlo experiment.

The empirical power for the $\mathcal{Z}_0$ test statistic in Proposition 1 is presented in Table 3. For illustration we also compare the power of our test to the power of the corresponding t-test statistic in Im, Pesaran, and Shin (2003), henceforth abbreviated to the IPS test. It should be noted that we compare the power to the IPS test that is based on a linear panel without a time trend.[7] This is reasoned by the fact that our panel lacks the property of a time trend. We also note that the same null hypothesis is tested.

In Table 3 we see that the power for the IPS test is rather poor for all T and n. Even for as large panels as $T = 100$ and $n = 50$, the power is only 0.21. The reason for the poor performance of the IPS test is that it is not designed to have power against nonlinear alternatives with a shift in levels. However, as indicated in He and Sandberg (2005a), the classical univariate Dickey-Fuller t-type of test also has poor power results (close to zero or equal to zero) in a similar set-up to (26). In our case this illuminates that a panel approach does not resurrect the

[7]The IPS test that we are using averages the classical Dickey-Fuller t-type of test based on an AR(1) process without a time trend.

Table 3: Empirical power of the $\mathcal{Z}_0$ test in Proposition 1 and the IPS test. The DGP is a PLSTAR(1) model. No serial correlation.

T	25		50		100	
n	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS
5	0.47	0.00	0.63	0.00	0.87	0.10
25	0.96	0.00	0.99	0.00	1.00	0.15
50	1.00	0.00	1.00	0.00	1.00	0.21

Note: The nominal size is 5%, and the results are based on 10 000 replications.

power for the Dickey-Fuller t-type of test in the univariate context to satisfactory levels. On the other hand, the $\mathcal{Z}_0$ test performs very satisfactorily and the power is close to unity for $n \geq 25$ and all samples sizes. As a final remark, the power results for the $\mathcal{Z}_0$ test in Table 3 is lower than for the test statistic used in He and Sandberg (2005c) under a similar Monte Carlo experiment (cf. Table 3 in He and Sandberg (2005c)). This might be explained by that the test statistic used in He and Sandberg (2005c) is of deviation type (the deviation form of the LSDV estimator) and not of t-type, and is based on pooled data. These two facts contributes to that the power for the $\mathcal{Z}_0$ test is lower.

5.2.2 A nearly linear heterogeneous panel

In the fourth Monte Carlo experiment we investigate the empirical power properties for $\mathcal{Z}_0$ when $\gamma_i = 0.01$ for all i in (26), and the resulting model in (25) is nearly linear. The set-up for the remaining parameters are the same as in (26). The results are presented in Table 4.

Table 4: Empirical power of the $\mathcal{Z}_0$ test in Proposition 1 and the IPS test. The DGP is an almost linear PLSTAR(1) model. No serial correlation.

T	25		50		100	
n	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS
5	0.12	0.96	0.46	1.00	1.00	1.00
25	0.46	1.00	1.00	1.00	1.00	1.00
50	0.76	1.00	1.00	1.00	1.00	1.00

Note: The nominal size is 5%, and the results are based on 10 000 replications.

From Table 4 we conclude that the power of the IPS test is superior, and is expected because the DGP is almost linear. Notable is, however, that the $\mathcal{Z}_0$ test statistic performs satisfactorily and for $n \geq 25$ and $T \geq 50$ the power equals unity. A nearly linear LSTAR(1) model is depicted in Figure 1.

5.2.3 A homogeneous nonlinear panel

In the fifth Monte Carlo experiment we abandon the randomness of the parameters in (25), and all cross sections are represented by the same LSTAR(1) model, i.e. a nonlinear homogeneous panel, with parameters

$$\pi_{i10} = 0, \quad \pi_{i11} = 0.1, \quad \gamma_i = 1, \quad c_i = 0.5T, \quad \pi_{i20} = 1, \quad \pi_{i21} = 0.8, \tag{27}$$

for all i. This scenario facilitates the comparison to the power of a similar univariate unit root test (parameter constancy test) in He and Sandberg (2005a) with the same DGP as in (25) and parameter values given by (27). The outcome of the experiment is presented in Table 5.

Table 5: Empirical power of the $\mathcal{Z}_0$ test in Proposition 1 and the IPS test. A nonlinear homogeneous panel. No serial correlation.

T	25		50		100	
n	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS
1*	0.11	0.00	0.20	0.00	0.28	0.00
5	0.67	0.00	0.76	0.00	0.90	0.00
25	1.00	0.00	1.00	0.00	1.00	0.00
50	1.00	0.00	1.00	0.00	1.00	0.01

Notes: The nominal size is 5%, and the results are based on 10 000 replications. The asterisk indicates that the power is calculated using the corresponding t-type of tests in Chapter 1 and in Dickey and Fuller (1979).

In Table 5 the benefits by using the panel unit root test statistic $\mathcal{Z}_0$ over a univariate t-type of tests are revealed. For an example, with $n = 1$ and $T = 100$ we see that the power is 0.28, and for $n = 5$ and $T = 100$ the power has increased to 0.90. A small increase in the number of cross sections leads to a substantial improvement in the power. Notable is also that the power for the IPS test is zero for all combinations of T and n. This should be compared to the results for the IPS test in Table 3, where the IPS showed some power, e.g. 0.21 for $T = 100$ and $n = 50$. The reason for the reduction in power for the IPS test is that the model in (27) generates time series with a more evident change in levels than the time series generated under the third Monte Carlo experiment, a fact that is illustrated in Figure 1. Even though a large amount of information is implied by a panel with dimensions $T = 100$ and $n = 50$, the panel is not, evidently, large enough.

5.2.4 A heterogeneous nonlinear panel with serially correlated errors

The sixth Monte Carlo experiment is conducted to examine the empirical power of the $\mathcal{Z}_3$ test statistic in Proposition 4. The DGP is the same as in (25) with parameter specifications as in (26), but the error process allows for serial correlation and is given by $u_{it} = 0.3u_{i,t-1} + v_{it}$, where $v_{it} \sim nid(0,1)$. The results for the sixth Monte Carlo experiment are presented in Table 6.

Table 6: Empirical power of the $\mathcal{Z}_3$ test in Proposition 4. The DGP is a PLSTAR(1) model. Serial correlation.

T	25			50			100		
	$\mathcal{Z}_3$			$\mathcal{Z}_3$			$\mathcal{Z}_3$		
n / p_i	1	2	3	1	2	3	1	2	3
5	0.11	0.10	0.09	0.25	0.21	0.19	0.75	0.73	0.71
25	0.28	0.25	0.23	0.41	0.35	0.31	0.95	0.94	0.92
50	0.42	0.38	0.35	0.58	0.51	0.49	1.00	1.00	1.00

Note: The nominal size is 5%, and the results are based on 10 000 replications.

When the errors are serially correlated, we see that the power for the $\mathcal{Z}_3$ test is reduced compared to the power results for $\mathcal{Z}_0$ with serially uncorrelated errors in Table 3. This is partially explained by that in the present case more parameters are estimated. The discrepancy in power between the $\mathcal{Z}_0$ and $\mathcal{Z}_3$ tests is therefore more pronounced for small T, and when increasing T the difference in power between the two tests become modest. Furthermore, ignoring the serial correlation (i.e. the case $p_i = 0$) yields an alarming situation, however not reported here, because the power is then close to zero for all combinations of T and n. Finally note that for $T = 100$, all n, and $p_i \geq 1$, the power is satisfactory.

5.2.5 A homogeneous panel with a mixture of nonlinear and non-stationary models

In the seventh and the last Monte Carlo experiment we examine the power when the DGP is a mixture of nonlinear models and unit root processes. The design of the experiments is the following: The nonlinear model is a PLSTAR(1) model with parameters given by (27), and the panel unit root processes are given in (21). Furthermore, let n_1 and n_2 denote the number of nonlinear models and unit root processes, respectively, such that $n_1 + n_2 = n$ holds. The simulation results are given in Tables 7.

Table 7: Empirical power of the $\mathcal{Z}_0$ test in Proposition 1 and the IPS test. The DGP is a mixture of LSTAR(1) models and unit root processes.

$n = 25$	(n_1, n_2)											
	$(25, 0)$		$(20, 5)$		$(15, 10)$		$(10, 15)$		$(5, 20)$		$(0, 25)$	
T	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS	$\mathcal{Z}_0$	IPS
25	1.00	0.00	0.99	0.00	0.88	0.00	0.62	0.00	0.20	0.00	0.05	0.05
50	1.00	0.00	1.00	0.00	0.90	0.00	0.64	0.00	0.23	0.00	0.05	0.05
100	1.00	0.01	1.00	0.00	0.96	0.00	0.66	0.00	0.27	0.01	0.05	0.05
250	1.00	0.29	1.00	0.19	1.00	0.16	0.92	0.14	0.59	0.12	0.05	0.05

Note: The nominal size is 5%, and the results are based on 10 000 replications.

From Table 7 we conclude that the power of the $\mathcal{Z}_0$ test is decreasing in n_2 towards the nominal size obtained with $n_2 = 25$, cf. also Table 1. On the other hand, the power of the $\mathcal{Z}_0$ test is increasing in n_1 and with $n_1 = 25$ the same results as in Table 5 are obtained. The power for the $\mathcal{Z}_0$ test is also increasing in T for all fractional combinations of nonlinear and unit root process. Furthermore, it is seen that when a small to medium fraction of the panel is non-stationary, the power of the $\mathcal{Z}_0$ test is barely affected and is close to unity. The power for the IPS test is close to zero because the panel is a mixture of nonlinear and non-stationary processes, and one can not expect the IPS test to gain any power by varying the fraction between these two options.

As a final remark, the results in Table 7 also illuminate the important aspect of that the outcome of panel unit root tests should be conservatively dealt with, meaning that rejecting the null hypothesis does not imply that all cross sections are nonlinear. For instance, in the case $n_1 = 5$, $n_2 = 20$, and $T = 250$, the power of the $\mathcal{Z}_0$ test equals 0.59, which means that we might erroneously model the whole panel as nonlinear when in fact only a small fraction of the panel is nonlinear. This also indicates that a careful joint analysis of both the individual and the panel unit root test results is called for to fully assess the stationarity/nonlinearity properties of the panel data, as pointed out in Karlsson and Löthgren (2000).

6 Concluding remarks

In this paper we present a new test for unit roots in a nonlinear dynamic heterogeneous panel. Theirs necessity can be motivated by the fact that canonical panel unit root test, such as the tests in Im, Pesaran, and Shin (2003), do not have satisfactory power when the DGP is an model with shifts in levels. Recent research indicates that many single time series exhibit a nonlinear adjustment path (structural shifts in levels) towards a long-run equilibrium. It seems therefore natural that the cross sections in a panel are modelled in a nonlinear way (or at least a fraction of them) as well.

Our nonlinear dynamic heterogeneous panel is general in the sense that it nests the PLSTAR model in He and Sandberg (2005c), a panel threshold autoregressive model, as well as the linear autoregressive panel in Im, Pesaran, and Shin (2003). Our panel is characterized by the fact that each cross section is modelled as an LSTAR model and where all parameters are assumed to be cross section specific. The residuals in the PLSTAR model are specified to be independent (a strong assumption) with heterogeneous variance among cross sections, but possibly serially correlated over time.

Our test for a common unit root in a nonlinear panel is based on averaging all individual t-statistics of a unit root for a specific cross section unit. The test is derived under: (i) T and n fixed. (ii) T fixed and letting $n \to \infty$. (iii) First letting $T \to \infty$ and thereafter letting $n \to \infty$ (sequential asymptotics). In the two latter cases it is shown that the limiting distribution of the test is the standard normal distribution.

Monte Carlo studies are performed and it is shown that the size distortions are negligible when the errors are uncorrelated over time. However, when the errors are modelled as an AR(1) process, it is crucial to include sufficiently many lags of the difference of the dependent variable, otherwise the test will be undersized. The power results are very satisfactory and are close to unity for panels bothbeing nearly linear or displaying a smooth shifts in levels and dynamics, as long as $n \geq 25$ and $T \geq 50$. In contrast, the IPS test only has power when a linear panel is considered, and otherwise its power is inferior. The improvement in power compared to a univariate test in He and Sandberg (2005a), is pronounced.

Appendix A: Simulated moments

Table 8 presents the simulated expected values, variances and third moments for the t_i^m test statistics in (8). For comparison, the same moments for the t_i test statistic in (7) are calculated as well. The moments are calculated under the null hypothesis in (6) assuming that the error term fulfills Assumption 1 with $\sigma_i^2 = 1$ for all i. It is seen that, for all T, $\tilde{S}_i^2$ is a very accurate estimate of σ_i^2, whereas S_i^2 is biased downwards. In Table 8 it is also revealed that the well-known fact of the LS bias in non-stationary regressions, see for instance Abadir (1993), is present in our case as well. Furthermore, we conclude that $\lim_{T\to\infty} \mu(T) = \mu^* = -1$ and $\lim_{T\to\infty} \eta^2(T) = \eta^{2*} = 1.43$, where $\mu(T) = E[t_i^m]$ and $\eta^2(T) = V[t_i^m]$, which confirms that the first two asymptotic moments exist, justifying the use of the Lindberg-Lèvy central limit theorem in Proposition 3.

Table 8: Simulated moments for t_i and t_i^m.

T	$E[t_i^m]$	$V[t_i^m]$	$E\,(t_i^m)^3$	$\tilde{S}_i^2$	$E[t_i]$	$V[t_i]$	$E\,(t_i)^3$	S_i^2
10	-0.540	0.898	-1.40	0.997	-0.716	2.086	-4.81	0.649
25	-0.786	1.160	-2.92	0.997	-0.901	1.583	-4.82	0.827
50	-0.889	1.278	-3.84	0.997	-0.950	1.502	-4.95	0.907
100	-0.940	1.361	-4.36	0.998	-0.972	1.466	-4.96	0.951
250	-0.974	1.401	-4.75	0.998	-0.992	1.444	-4.97	0.980
500	-0.985	1.427	-4.87	0.998	-0.995	1.442	-4.98	0.989
1000	-0.991	1.430	-4.97	0.999	-0.997	1.438	-5.00	0.995
∞	-1.000	1.432	-5.03	1.000	-1.000	1.432	-5.03	1.000

Note: The results are based on 1 000 000 replications.

The simulation results for the expected values and variances for t_i^a are reported in Tables 9 and 10 respectively. It is pointed out that both the expected value and the variance for t_i^a depends on the factors T, $\mathbf{y}_{i0}$, p_i, $\boldsymbol{\zeta}_i$, and σ_i^2. In order to present these moments, the simulation results are given for various combination of T and p_i when they are both considered as fixed and known. We let $\mathbf{y}_{i0} = \mathbf{0}$ to capture the invariance with respect to σ_i^2. However, the problem with the nuisance parameters $\boldsymbol{\zeta}_i$ under the null hypothesis still remains. The option we choose is to set $\boldsymbol{\zeta}_i = 0$ under the null hypothesis so that data are generated from the model $y_{it} = y_{i,t-1} + u_{it}$ where $u_t \sim nid(0,1)$. On these data we run the regression in (18) and calculate the t_i^a statistic in (22). This is repeated 1 000 000 times, for each desired sample size T, to generate the expected values and variances for t_i^a. This means that we are approximating e.g. $E[t_i^a(T, p_i, \boldsymbol{\zeta}_i)]$ with $E[t_i^a(T, p_i, \mathbf{0})]$ under the null hypothesis. Without presenting the results we note that these expected values and variances are rather robust against the values of $\boldsymbol{\zeta}_i$, and that the approximations for the expected values and the variance are reasonable.

Table 9: Simulated expected values for t_i^a.

$E[t_i^a(T, p_i, \boldsymbol{\zeta}_i = \mathbf{0})]$

p_i / T	1 ζ_{1i}=0	2 ζ_{1i}= ζ_{2i} =0	3 $\zeta_{1i} = \zeta_{2i} = \zeta_{3i}$=0
25	-0.993	-0.993	-1.052
50	-0.994	-1.000	-1.049
100	-1.000	-1.000	-1.034
250	-1.000	-1.000	-1.011
500	-1.000	-1.000	-1.000
1000	-1.000	-1.000	-1.000
∞	-1.000	-1.000	-1.000

Note: The results are based on 1 000 000 replications.

Table 10: Simulated variances for t_i^a.

$V[t_i^a(T, p_i, \boldsymbol{\zeta}_i = \mathbf{0})]$

p_i / T	1 $\zeta_{1i} = 0$	2 $\zeta_{1i} = \zeta_{2i} = 0$	3 $\zeta_{1i} = \zeta_{2i} = \zeta_{3i} = 0$
25	1.661	1.713	1.762
50	1.553	1.585	1.611
100	1.495	1.513	1.535
250	1.473	1.462	1.484
500	1.451	1.444	1.452
1000	1.441	1.435	1.434
∞	1.432	1.432	1.432

Note: The results are based on 1 000 000 replications.

For $T \to \infty$ in Tables 9 and 10 we see that both the expected value and the variance for t_i^a equal the expected value (i.e. μ^*) and the variance (i.e. η^{2*}) for t_i^m in Table 8. This is an implication of Corollary 2.7 in He and Sandberg (2005b).

References

Abadir, K. M. (1993): "OLS bias in a nonstationary autoregression," *Econometric Theory*, 9, 81–93.

Dickey, D. A., and W. A. Fuller (1979): "Distribution of the Estimators for Autoregressive Time Series," *Journal of the American Statistical Association*, 74, 427–431.

Harris, R. D. F., and E. Tzavalis (1999): "Inference for unit roots in dynamic panels where the time dimension is fixed," *Journal of Econometrics*, 91, 201–226.

Harvey, D. I., and T. C. Mills (2002): "Unit Roots and Double Smooth Transitions," *Journal of Applied Statistics*, 29, 675–683.

He, C., and R. Sandberg (2005a): "Testing Parameter Constancy in Unit Root Autoregressive Models Against Continuous Change," Stockholm School of Economics, SSE/EFI Working Paper Series in Economics and Finance No. 579.

——— (2005b): "Dickey-Fuller Type of Tests Against Non-Linear Dynamic Models," Stockholm School of Economics, SSE/EFI Working Paper Series in Economics and Finance No. 580.

——— (2005c): "Inference for unit roots in a panel smooth transition autoregressive model where the time dimension is fixed," Stockholm School of Economics, SSE/EFI Working Paper Series in Economics and Finance No. 581.

Hsiao, C. (1986): *Analysis of panel data*. Cambridge University Press.

Im, K. S., M. H. Pesaran, and Y. Shin (2003): "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 115, 53–74.

Karlsson, S., and M. Löthgren (2000): "On the power interpretation of panel unit root tests," *Economic Letters*, 66, 249–255.

Lanne, M., H. Lütkepohl, and P. Saikkonen (2003): "Test procedures for unit roots in time series with level shifts at unknown time," *Oxford Bulletin of Economics and Statistics*, 65, 91–115.

Larsson, R. (1997): "On the asymptotic expectations of some unit root tests in a first order autoregressive process in the presence of a trend," *Annals of the Institute of Statistical Mathematics*, 49, 585–599.

LEVIN, A., C. LIN, AND C. CHU (2002): “Unit root tests in panel data: Asymptotic and finite-sample properties,” *Journal of Econometrics*, 108, 1–24.

LEYBOURNE, S., P. NEWBOLD, AND D. VOUGAS (1998): “Unit roots and smooth transitions,” *Journal of Time Series Analysis*, 19, 83–97.

LIN, C. F. J., AND T. TERÄSVIRTA (1994): “Testing the constancy of regression parameters against continuous structural change,” *Journal of Econometrics*, 62, 211–228.

LUUKKONEN, R. P., P. SAIKKONEN, AND T. TERÄSVIRTA (1988): “Testing linearity against smooth transition autoregressive models,” *Biometrika*, 75, 491–499.

MADDALA, G. S., AND S. WU (1999): “A Comparative Study of Unit Root Tests With Panel Data and a New Simple Test,” *Oxford Bulletin of Economics and Statistics*, pp. 631–652.

NABEYA, S. (1999): “Asymptotic moments of some unit root test statistics in the null case,” *Econometric Theory*, 15, 139–149.

PESERAN, M. H. (2003): “A simple panel unit root test in the presence of cross section dependence,” Cambridge University, DAE Working Paper No. 0346.

PHILLIPS, P. C. B., AND H. MOON (1999): “Linear regression limit theory for nonstationary panel data,” *Econometrica*, 67, 1057–1111.

PHILLIPS, P. C. B., AND D. SUL (2003): “Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence,” Cowles Foundation Discussion Paper No. 1438.

QUAH, D. (1994): “Exploiting cross section variation for unit root inference in dynamic data,” *Economic Letters*, 44, 9–19.